Journal of Risk Research **2** (1), 11–29 (1999)

Prioritizing stakeholder concerns in environmental risk management

ROBERTO ACCORSI, GEORGE APOSTOLAKIS* AND ENRICO ZIO‡

Department of Nuclear Engineering, Room 24–221, Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA

Abstract

The management of environmental programmes often involves several stakeholders with diverse, and often conflicting, concerns. This paper proposes a methodology for the assessment of stakeholder preferences regarding a number of objectives of environmental restoration activities, such as the minimization of costs and of the impact on human health and safety. This methodology is based on an analytic/deliberative process that starts with pairwise comparisons of these objectives using questionnaires that the stakeholders complete. This input is used to produce a first set of relative weights using the Analytic Hierarchy Process (the analytic part). This set, as well as possible inconsistencies of the stakeholder assessments, is discussed with the stakeholders to correct inconsistencies and revise the weights (the deliberative part). The stakeholders always have the final word regarding the relative weights. Insights gained from a case study are also presented. The feedback from the stakeholders participating in this exercise was positive.

1. Introduction

Environmental management problems involve a variety of concerns, such as cost, health and programmatic risks, as well as impacts on cultural resources. A problem with the analysis of these impacts of environmental decisions is that they have different units and are difficult to define. Yet, they must be accounted for in an efficient decision-making process, as failure to address any of these is likely to result in controversial decisions.

A second problem is due to the fact that there are usually several individuals or groups that are interested in influencing the decision, even though they are not the decision-makers themselves. These individuals or groups are generally referred to as stakeholders and they may include various government agencies and citizen groups. It is widely accepted that the stakeholders must be involved in the decision-making process from the beginning (National Research Council, 1996). Consequently, the decision-maker must deal with a diverse audience of technical and non-technical people, as well as with the multitude of their technical and value-laden quantitative and qualitative concerns. Each stakeholder has his/her own set of objectives that he/she wishes to

*To whom correspondence should be addressed.
‡Permanent address: Politecnico di Milano, Dipartimento di Ingegneria Nucleare, via Ponzio 34/3, 20133 Milano, Italy.

achieve. Although other stakeholders share most of these objectives, they may not be valued the same way. For example, the costs of the proposed decision alternatives are usually viewed differently by the site owner and the other stakeholders. It is evident, therefore, that methods are needed to elicit stakeholder priorities and preferences that are accessible to all kinds of stakeholders and flexible enough to account for the imprecision and inconsistencies inherent in human subjective judgements.

The decision-making process cannot be based on the prioritization of the decision criteria only. The alternatives at hand must be analysed and then compared against these criteria to see which alternative is most likely to match them. It is unlikely that, in a real case, an alternative will perform better than all the others under all criteria. For instance, inexpensive alternatives typically result in a relatively poor removal of contamination or may not be as safe for workers as more expensive ones. Therefore the decision making process is seldom cut and dry and one should consider some trade-offs, for instance, whether improvements are worth their cost. This depends on the actual extent of the improvement, its cost and on how these are valued by different stakeholders.

Thus, two sets of data must be combined: the performance of the decision alternatives under different criteria (such as cost, contamination removal, and worker safety) and the relative importance that the stakeholders wish to assign to the decision criteria. Once these two elements are combined, an overall Performance Index (*PI*) can be defined for each alternative for a given stakeholder, who can, then, rank the decision alternatives.

This integration process can be carried out using the following equation, whose elements summarize this discussion:

$$PI_j^k = \sum_i w_i^k u_{ij}^k \tag{1}$$

where PI_j^k is the performance index for the jth decision alternative, for the kth stakeholder; w_i^k is the weight of the ith decision criterion (or Performance Measure, elicited from the kth stakeholder, see Section 2.1); u_{ij}^k is the utility of the impact of the jth decision alternative on the ith decision criterion (also elicited from the kth stakeholder; the impacts are estimated by risk assessments).

This paper presents an application of the Analytic Hierarchy Process (AHP; Saaty, 1996) to the problem of developing a set of weights for each stakeholder that capture his/her preferences over a set of decision criteria. In other words, the w_i^k are being assessed. This development is conducted in an interactive way in which the AHP results are the starting point and the stakeholder, through deliberation, has the final word.

This methodology was implemented in a project whose aim was to develop a decision-making framework involving multiple stakeholders that could be used by the Department of Energy in environmental restoration decisions (Advanced Sciences, Inc., Beta Corporation International, Massachusetts Institute of Technology, New Mexico State University, and QuanticSci Inc., 1997). The present paper deals only with the derivation of weights. The site chosen for the development of this prototypical methodology was a hazardous chemical waste landfill (CWL) located at a US National Laboratory. To simplify the analysis, the project team focused on TCE (trichloroethylene, an organic) and chromium (a metal), which were considered to be the primary environmental concerns at the site. Eleven stakeholders representing the public, the site owner, and regulatory agencies (state and city) participated in the prioritization (see Table 1). The decision options consisted of six Remedial Action Alternatives (RAAs) ranging from in-situ vitrification to off-site treatment.

It should be stressed that the assessment of the impacts (u_{ij}^k) and the computation of the PI_j^k are outside the scope of this work. Also, for an accurate description of how the *PI* rankings for different stakeholders were used on the way towards a decision in the context of an analytic-deliberative process (National Research Council, 1996), the reader is referred to Apostolakis and Pickett (1998).

In Section 2, we present the structure of the decision problem and the overall methodology for the derivation of weights. The aim is to provide some more detail on Equation 1. In Section 3, we review the Analytic Hierarchy Process, which is the method for pairwise comparisons of objectives that has been used in this work. The actual implementation of the proposed approach is presented in Section 4, along with several insights gained during this application. Finally, Section 5 offers some conclusions.

2. Structure of the decision problem

In this section we outline a systematic process of structuring the decision problem by identifying and decomposing the concerns and objectives of the stakeholders.

2.1. Performance measures

The first step is the construction of a hierarchical structure of impacts and objectives often called the *value tree*. Examples of how this can be done are given in Edwards and Newman (1986) and in Gregory and Keeney (1994). Although the development of this tree is beyond the scope of this paper, we discuss its parts that will help us in explaining the prioritization process.

In our case study, the top objective is to 'Maximize the Benefits of Remediation'. The question now is how each RAA contributes to the achievement of this general objective. For the CWL, the analysts identified, in close collaboration with the stakeholders, six broad categories of RAA impacts: 'Programmatic Assumptions', 'Life Cycle Cost', 'Socioeconomic Issues', 'Cultural, Archaeological, and Historic Resources', 'Environment' and 'Human Health and Safety' (Fig. 1). An alternative way of looking

Table 1. Stakeholder affiliations.

Stakeholder	*Organization*
SH1	Real estate agent
SH2	National laboratory employee
SH3	City Environment and Health Department
SH4	Regional Council of Governments
SH5	National laboratory contractor
SH6	Community Advisory Board
SH7	Community Involvement and Issues Management
SH8	Native American Nations
SH9	State Environmental Department
SH10	City
SH11	County Environment and Health Department

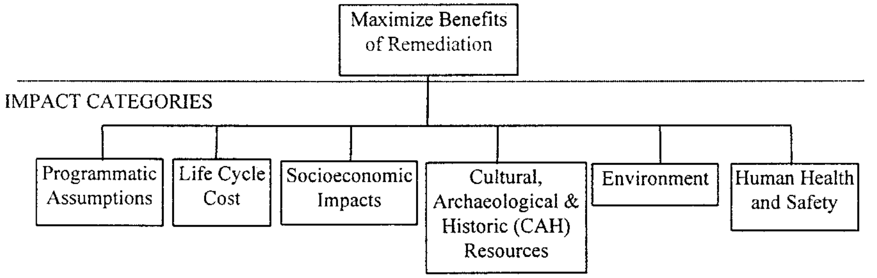

Fig. 1. First tier of the value tree: impact categories.

at these impacts is to say that the minimization of, for example, the impact on human health and safety is an objective category that contributes to the maximization of remediation benefits, i.e. the overall goal.

These impact categories are very broad and do not lend themselves to a direct quantitative evaluation. The categories, then, are decomposed into their more elementary constituents; in other words, one should define which particular objectives make up each impact category. The objectives represent the specific action goals that the decision should achieve. Within the 'Socioeconomic Impacts' category, for example, 'Promoting Community Quality of Life' and 'Environmental Justice' are two objectives that the implementation of an RAA should achieve (Fig. 2). As a second example, an RAA can be preferred over another because it is easier to implement (e.g. requires fewer mechanical equipment) or results in minimal waste generation or, finally, presents fewer organizational concerns (e.g. it does not need any permission for the use of special technologies or is less likely to be subject to administrative delays). This would be captured

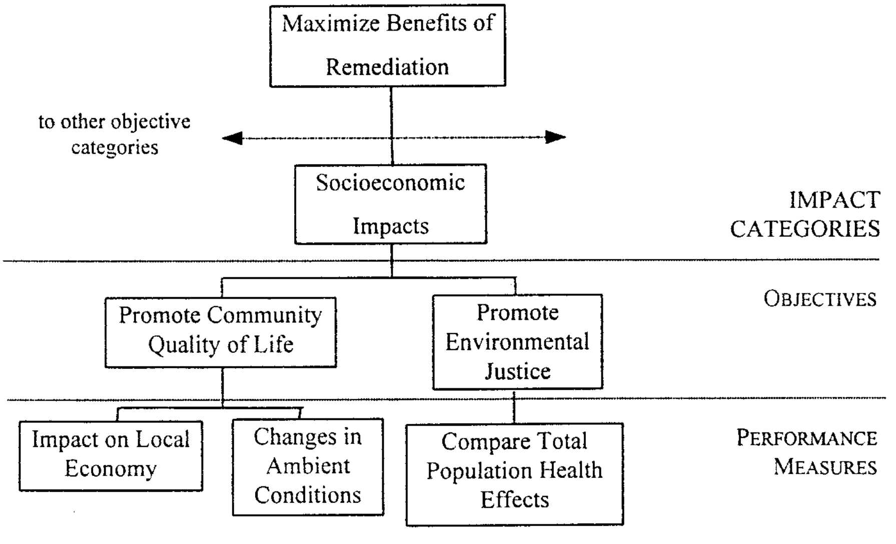

Fig. 2. Socioeconomic impact category broken into its objectives and performance measures.

by the objectives constituting the category 'Programmatic Assumptions'. Similarly, all the objective categories can be expanded in terms of the related objectives.

The objectives are further refined in terms of Performance Measures. These are the measurable quantities needed to evaluate how well an RAA meets the objectives (they are the 'decision criteria' of Equation 1). For example, the impact that a remediation technology has on the local economy and the changes it introduces to the ambient conditions can be considered directly quantifiable measures of how a given RAA promotes the community quality of life (Fig. 2). The transition from a general level (the impact categories) to a more specific, and quantifiable, level (the PMs) is now complete (Fig. 3).

The result of the structuring process is not expected to be unique in the sense that different stakeholders are expected to organize their objectives differently. Even in the case of a tree common to a number of stakeholders, the priority weights assigned to the branches of the tree, being a measure of the relative importance of the corresponding issues, are expected to be different for different stakeholders.[1]

2.2. Decision analysis

Formal decision theory can be applied to each stakeholder's preferences to develop a ranking of his/her decision alternatives. Examples of applications of decision theory to environmental problems can be found in Merkhofer and Keeney (1987), Keeney and von Winterfeldt (1994), and Kadvany (1995). In general, this requires the evaluation of the impact of each remedial action alternative on the set of Performance Measures and the expression by the stakeholder of his/her preferences over these impacts. This allows the calculation of the expected utility for each alternative. To distinguish the latter from the expected utility of each PM (u_{ij}^k in Equation 1), we call the expected utility of a decision alternative its Performance Index (*PI*; see Equation 1).

Behind the formal structure of Equation 1 there is an assumption of mutual preferential independence of the performance measures (Keeney, 1981) which needs to be checked with the stakeholders. It must be stressed though, that linear models are rather robust, so that even quite substantial amounts of deviation from the independence assumption will make little difference to the ultimate utility values and even less to their ordering.

To obtain the quantification of the impact of each RAA on the PMs, i.e., the u_{ij}^k, influence diagrams can be constructed (Hong and Apostolakis, 1993) and used with risk assessment codes. These implement mathematical models of both deterministic and stochastic phenomena. Their ultimate output is a set of values for the performance measures, whose distribution can, then, be analysed. A detailed discussion of these diagrams and the derivation of utility values are beyond the scope of this paper.

The performance indices allow each stakeholder to rank the available decision alternatives in order of preference (the most desirable alternative has the highest *PI*). These results are the starting point of a deliberative process that produces the ultimate recommendations from the stakeholders to the decision maker (Apostolakis and Pickett, 1998).

[1]In our case study, some stakeholders felt that the objective 'Minimize Long-Term Risk to Public Health and Safety' should be in the category 'Human Health and Safety' (Fig. 3). Most of them, however, wished to move this objective under the category 'Environment' and to change the name of the category 'Human Health and Safety' to 'Worker Health and Safety.' Accordingly, two value trees, were developed and analysed.

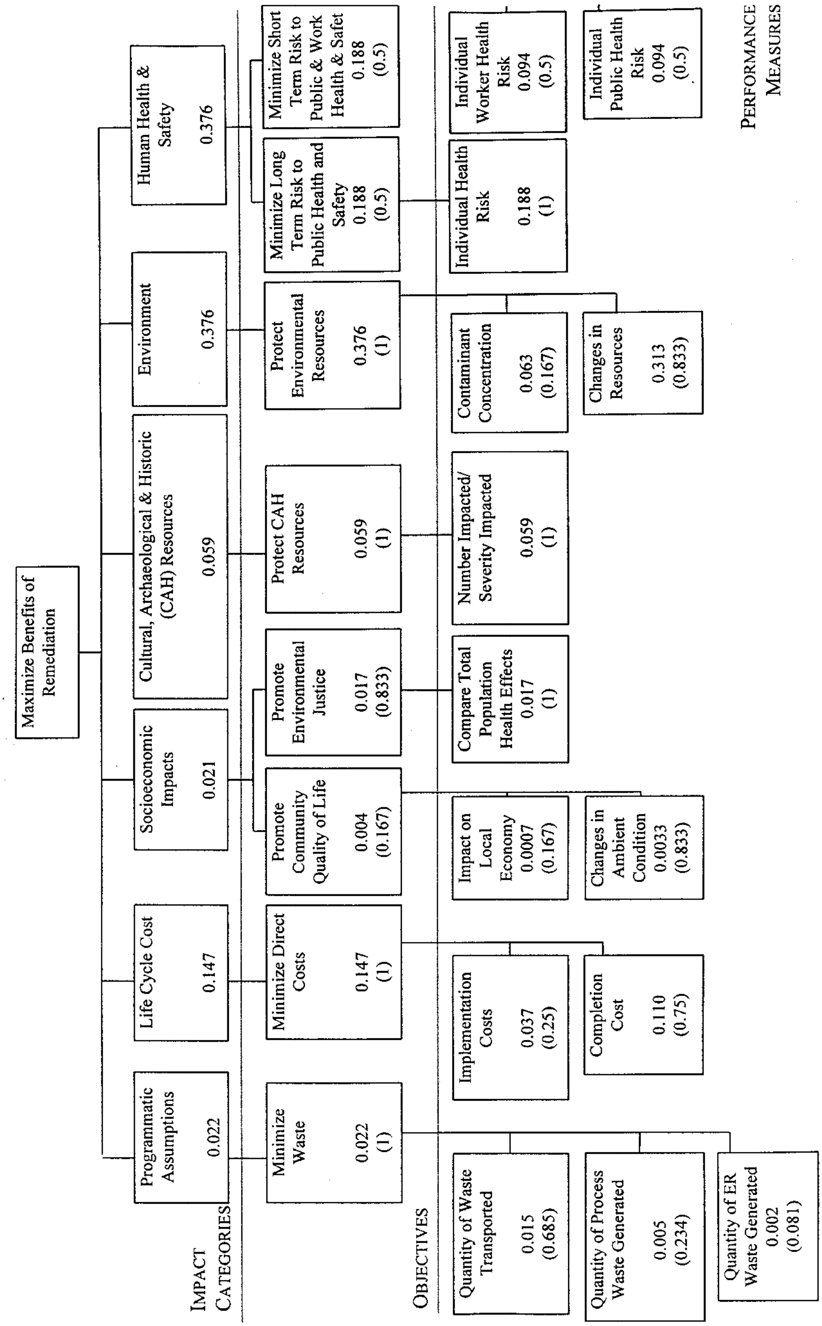

Fig. 3. Example of complete prioritization (SH2).

2.3. Combination of weights

As it is evident from Equation 1, the weights needed for making a decision are the weights of the performance measures themselves. Therefore, strictly speaking, one could skip the assessment of higher-level weights, e.g., at the impact category level, and compare directly the PMs with each other. This, however, is not easy for the stakeholders. Comparing elements at the same level in the tree, e.g., impact categories first, then objectives within each category, has been found to be much easier for people to handle. It also reduces the total number of comparisons involved.

The procedure starts at the impact category tier: all the categories are given a weight, ranging from 0 to 1, so that the ratios reflect the relative importance of the categories and the sum of all the weights is unity. This is accomplished through pairwise comparisons, as we will discuss in the next section.

The second step consists of the comparison of the objectives within each category. As done at the higher level, the objectives are compared two at a time and the result is a weight between 0 and 1 for each objective. Again, the weights are normalized to unity. It is important to note that the weights so obtained reflect the relative importance of the objectives *within* the impact category. This implies that, at this stage, it would be meaningless to compare the weights of objectives that belong to different categories. This can be done only after multiplying each objective's weight by the weight of the corresponding category.

As an example, Fig. 3 shows that, using input from stakeholder no. 2, the objective 'Promote Environmental Justice' is assigned a weight of 0.833, while the objective 'Promote Community Quality of Life' is assigned a weight of 0.167 (both are under the category 'Socioeconomic Impacts'). At the same time, under 'Human Health and Safety', the objective 'Minimize Long Term Risk to Public Health and Safety' was judged as important as the objective 'Minimize Short Term Risk to Public & Worker Health & Safety'. Each objective received a relative weight of 0.50. This does not mean, however, that the objective 'Promote Environmental Justice' is this stakeholder's main concern. In fact, the category 'Human Health and Safety' is more important than the category 'Socioeconomic Impacts' (0.376 versus 0.021). The absolute weights of the objectives are obtained by multiplying the relative weights (i.e., within each category) by the weight of the category itself. Thus, the objectives 'Minimize Long Term Risk to Public Health and Safety' and 'Minimize Short-Term Risk to Public & Worker Health & Safety' receive a weight of 0.188 (0.50×0.376) each, while the objective 'Promote Environmental Justice' receives 0.017 (0.833×0.021) only, reflecting this stakeholder's concern for Health and Safety.

The last step is the comparison of PMs relative to the same objective. When the weights are multiplied by the absolute weights of the objectives, absolute weights are obtained for the PMs. The results for stakeholder no. 2 are shown in Fig. 3.

It is important to emphasize that the pairwise comparisons must be carried out in the context of the specific problem under investigation. When the stakeholders were asked to perform pairwise comparisons at the impact category level (for example), they were not comparing 'Human Health and Safety' to 'Socioeconomic Impacts' in general, but, rather, for the specific problem at hand. This required a discussion with the stakeholders of the ranges of impacts that were expected for this problem. As the actual performance measure ranges had not been fully developed yet in our case study, this discussion was largely qualitative with some quantitative estimates. In the case of this site, the contamination was not high and the stakeholders were fully aware of the fact that no severe impacts were anticipated. When the actual ranges became available several weeks later, one stakeholder who had been particularly interested in the ranges was asked whether she would like to modify the weights that she had assigned. She declined stating that the original set was representative of her preferences even in the light of the new quantitative information.

Finally, it should be once more emphasized that the approach addresses only the individual preference of the stakeholder, and it does not provide any direct quantitative or qualitative method to reconcile the possible conflicts among stakeholders. On the other hand, the approach is structured in such a way that such possible conflicts are easily traceable down to their basic constituents and reasons, so that many insights can be provided to the decision-maker who has to perform the integration of the various stakeholders' preferences. The deliberative process that integrated the stakeholders' views is described in Apostolakis and Pickett (1998).

The problem is now reduced to the elementary prioritization procedure, that is, the weight assessment within a given level. This involves pairwise comparisons, as we will discuss in the next section.

3. The Analytic Hierarchy Process

The weights for the elements on a given tier of the value tree could be obtained by simply asking the stakeholders to assign them. This apparently simple method, however, is not satisfactory because it does not delve deeper into the stakeholder preferences. This is important since it is well-known that people have difficulties expressing their preferences in a consistent manner. Thus, a method that allows us to elicit information at a more elementary level is desirable. At the same time, we recognize that such a method is bound to be fraught with uncertainties. Consequently, we have decided to use a structured method for preference elicitation and, then, to have a session with each stakeholder in which the results of this method are discussed, inconsistencies are pointed out, and the set of final weights is produced. In other words, we do not rely exclusively on the chosen method to produce the weights.

A number of methods for estimating the weights w_i^k has been proposed in the literature (Clemen, 1991). One approach that has been implemented in risk assessments is to estimate a monetary equivalent for each PM. An example of how monetary equivalents and constructed scales are formed can be found in Keeney and von Winterfeldt (1994), where strategies to manage nuclear waste from power plants are compared. A drawback of the method is the emotional response of the stakeholders when they are asked to express all PMs in terms of dollars, especially when delicate issues, such as the value of life, are involved.

The Analytic Hierarchy Process (AHP) recognizes that inconsistency in comparisons is expected and its investigation can lead to useful insights into the stakeholder's preferences.

To clarify this point, assume that three elements, say A, B and C are to be prioritized. This could be done by assessing the relative importance of A and B, and of B and C. In principle, one might say that if A is as important as B, and B is twice as important as C, then A must be twice as important as C. This reasoning, though, would rely on the hidden assumption that one can assess the importance ratio of two elements by comparing them indirectly with a third one. In our example, the relative preferences over A and C are obtained indirectly through comparisons with B. Strictly speaking, this information is indeed sufficient to supply all the weights, but unfortunately, in real cases, this seemingly obvious transitivity property – and the implied rational behaviour of the stakeholder – is not warranted. This means that, in a real case, if the apparently unnecessary direct comparison between A and C, is made, it is very likely that the result will be different from the expected one (i.e., the one obtained through indirect comparison with B). For this reason, more than the strictly necessary information is elicited.

In general, when n elements are compared, all the possible $[n(n-1)]/2$ pairwise comparisons are elicited (in the preceding example, n equals 3, therefore, there are three possible pairwise comparisons).

In AHP, judgements are ratio judgements, but are expressed in linguistic terms such as 'weakly preferred' or 'moderately preferred.' Table 2 shows the semantic scale. It is explained below.

Let us consider the prioritization of four elements A, B, C, and D. Pairwise comparisons are performed for all the possible $4 \times 3/2 = 6$ pairs. For each pair, the assessor is asked which element is more important and how strongly in terms of the semantic scale. Using Table 2, this assessment is converted to a number s that is designated as the entry a_{ij} of $\mathbf{A}$, the matrix of pairwise comparisons. The entry a_{ij} indicates by how much the ith element is preferred over the jth element. If the opposite is true, i.e., if the jth element is preferred over the ith element, the inverse of s is used for a_{ij}, i.e., $a_{ij} = 1/s$ and $a_{ji} = s$. Note that this does not assume any transitivity property, as no third element is involved. From this definition, it follows that the diagonal elements of matrix $\mathbf{A}$ are equal to unity, since any element is as important as itself.

Returning to our example, if the stakeholder estimates that C is somewhere between equally and weakly more important than D, then, from Table 2 we see that s is 2. C is more important, so s is inserted into the matrix at the 3rd row and 4th column (i.e., $a_{34} = 2$). The value of ½ is inserted in the symmetric position, i.e., $a_{43} = 1/2$. Similarly, all the elements of the matrix can be determined. Table 3 shows the matrix of pairwise comparisons for this hypothetical example. As an example, the entry a_{41} is equal to 1/8. This means that the assessor judged element A to be between demonstrably and absolutely more important than D. Since D has a rating of 1/8 with respect to A and C has a rating of 2 with respect to D, consistency requires that A have a rating of exactly 4 with respect to C, i.e., A must be somewhere between weakly and strongly more important than C. This would preserve the transitivity property among the

Table 2. Semantic scale used in an AHP prioritization.

1: if the two elements are **equally** important.
3: if one element is **weakly** more important than the other element.
5: if one element is **strongly** more important than the other element.
7: if one element is **demonstrably** or **very strongly** more important than the other element.
9: if one element is **absolutely** more important than the other element.
Numbers 2, 4, 6 and 8 can be used to express compromise between slightly differing judgments.

Table 3. Matrix **A** of pairwise comparisons and weight derivation via AHP.

Matrix **A**	A	B	C	D	*Weights*
A	1	2	4	8	0.53
B	1/2	1	2	4	0.27
C	1/4	1/2	1	2	0.13
D	1/8	1/4	1/2	1	0.07
				Total	1.00

elements A, C, and D. As the table shows, this is indeed the case here (Table 3 is constructed so that the judgements are consistent).

The priority weights are derived from the matrix **A**. A number of possible algorithms are discussed by Fichtner (1986). Saaty (1996) recommends computing the weights as the components of the eigenvector of **A** associated with the largest eigenvalue λ_{max} (the principal eigenvector), normalized so that the sum of its elements is unity. He offers an intuitive argument in support of what at first glance might appear to be an exotic method for weight calculation. A simple computer program carries out the necessary calculations. The last column of Table 3 contains the priority weights for the elements of A, B, C and D of the previous example (the principal eigenvector). Note that, in this case of consistent entries, all the columns are proportional to this eigenvector, while the rows are proportional to the inverse of its elements. In the inconsistent case, these properties do not hold.

Deviation from consistency is measured by the *inconsistency ratio*[2] which is defined as the ratio of the inconsistency index $[(\lambda_{max}-n) / (n-1)]$ to the expected value of the inconsistency index of an n by n matrix derived from random assessments (random index), where n is the number of elements being compared (and, thus, the dimension of the matrix).

In the case of perfect consistency λ_{max} is n (see footnote 2) and the index is zero. Accordingly, the inconsistency ratio is zero as well. On the other hand, if the comparisons are carried out randomly, the inconsistency ratio is 1. An inconsistency ratio of about 0.10 or less is usually considered acceptable. We note that this ratio is for the analyst's use and it is not communicated to the stakeholder. When the inconsistency is relatively high, the analyst looks for the inconsistencies in the stakeholder's pairwise comparisons and points them out to the stakeholder who is, then, free to change them so that greater consistency can be achieved. Telling the stakeholder that a mathematical quantity called the inconsistency ratio is high does not serve any purpose and, in fact, can be detrimental.

As stated earlier, after the calculations are carried out, it is imperative that the results be shown to the stakeholders, to give them the opportunity to confirm that they indeed reflect their judgement of the relative importance of the elements. This is of extreme importance, as it would be naive to assume that any procedure of this nature would be capable of producing the correct weights directly. Rather, the application of AHP

[2]The concept of inconsistency ratio is based on the observation that, in the consistent case, all rows are proportional to each other and all the eigenvalues are 0, except for the largest eigenvalue, which is equal to n. In mathematical terms, this is due to the fact that the matrix **A** has trace equal to n and rank equal to 1. Note that Saaty (1996) calls it the consistency ratio.

should be interpreted as a means of obtaining an approximate set of weights which is to be refined through successive iterations with the stakeholder. The inconsistency ratio can be used to identify the extent and root causes of potential inconsistencies thus focusing the discussion of the results on the points that have the highest potential of enhancing the accuracy of stakeholder input.

4. Insights gained from the case study

This section contains the insights gained from the case-study experience. For simplicity, the discussion is limited to the impact category level. As an example, however, Fig. 3 shows the relative weights (as obtained at each level, in parentheses) and the absolute weights for stakeholder no. 2. Similar trees were drawn for each stakeholder and there is no attempt to combine the trees or the weights assigned by the stakeholders at this point.

4.1. Interaction with the stakeholders

After the pairwise comparisons elicited from each stakeholder were processed to produce the relative weights, a session was held with each stakeholder separately to discuss his/her individual results. In general, most stakeholders agreed that the procedure had produced a correct ranking and sufficiently accurate weights. Some stakeholders, however, did raise questions about the derived results and requested further explanations from the analysts.

The representative from the County Environmental Health Department, stakeholder no. 11, provided the assessments summarized in Table 4. The calculated weights for the impact categories are reported in the same table. When he was shown these results, the stakeholder suggested that the category 'Life Cycle Cost' (LCC) should be ranked fourth, i.e., above the 'Socioeconomic Impacts' and 'Programmatic Assumptions' categories. The stakeholder added, however, that this was not a major concern to him, because the weights of these three categories were very close to each other.

We pointed out to the stakeholder that his suggestion was inconsistent with his earlier input. Indeed, as the matrix in Table 4 shows, in the direct comparison between 'Life Cycle Cost' and 'Programmatic Assumptions', the stakeholder had given priority to the former (the corresponding matrix entry, a_{65}, equals 3, indicating that this stakeholder considered LCC as weakly more important than 'Programmatic Assumptions'). This stakeholder's fairly high inconsistency ratio, 0.2, drew the analyst's attention and a more detailed scrutiny of the stakeholder's stated preferences was undertaken.

Considering the first four entries of the last two rows of the matrix in Table 4 indicates that the category 'Programmatic Assumptions' always dominates 'Life Cycle Cost'. In fact, a_{62} indicates that 'Life Cycle Cost' is judged to be three times less important than the category 'Cultural', which is judged as important as 'Programmatic Assumptions' (entry a_{52}). If the stakeholder were consistent, this would imply that the category 'Programmatic Assumptions' should be three times as important as 'Life Cycle Cost'. As this was repeated in every column (except, of course, for the direct comparison a_{65}), it is not surprising that the AHP ranked 'Life Cycle Cost' last. Thus, the stakeholder's desire to rank 'Life Cycle Cost' higher is not to be ascribed to a flaw in the algorithm, but to this stakeholder's inconsistency.

Having identified the problem with the last two rows, it is reasonable to investigate what happens when these inconsistencies are corrected. Thus, Table 5 is produced.

Table 4. Matrix of pairwise comparisons and weights for stakeholder no. 11.

	Environ	Cultural	HH&S	Socioec	Program	LCC
Environ	1	7	1/7	7	7	9
Cultural	1/7	1	1/5	3	1	3
HH&S	7	5	1	5	5	7
Socioec	1/7	1/3	1/5	1	1	3
Program	1/7	1	1/5	1	1	1/3
LCC	1/9	1/3	1/7	1/3	3	1

	HH&S	*Environ*	*Cultural*	*Socioec*	*Program*	*LCC*	*Total*
Weights	0.500	0.276	0.078	0.055	0.046	0.045	1

In this case, the inconsistency ratio is improved to 0.15. The ranking is now exactly the ranking that the stakeholder suggested.

Another interesting insight came from stakeholder no. 6, whose ranking was the one shown in Table 6(I). After reviewing this result, the stakeholder did not agree and directly expressed the preferences of Table 6(II). The major difference is the reversal in the ranking of the categories 'Programmatic Assumptions' and 'Cultural'. Although in the questionnaire itself the direct comparison was in favour of the 'Cultural' category, as shown in the weight vector of Table 6(I), the stakeholder was surprised that she had done so. She, then, proceeded to change the weights to the new set shown in Table 6(II). This is a good example of direct fine-tuning of the results by the stakeholder. The AHP analysis gives a first set of preferences, which the stakeholder is free to adjust after reconsideration of the original pairwise comparisons.

In the next example (unlike the preceding two examples), the interaction with the stakeholder was initiated by the analysts (i.e., us), after we noticed a high inconsistency ratio. Table 7 shows the results for stakeholder no. 1. The inconsistency ratio is 0.1930. Upon investigation, we realized that if the category 'Life Cycle Cost' were not considered,

Table 5. New matrix for stakeholder no. 11.

	Environ	Cultural	HH&S	Socioec	Program	LCC
Environ	1	7	1/7	7	9	7
Cultural	1/7	1	1/5	3	3	1
HH&S	7	5	1	5	7	5
Socioec	1/7	1/3	1/5	1	3	1
Program	1/9	1/3	1/7	1/3	1	1/3
LCC	1/7	1	1/5	1	3	1

	HH&S	*Environ*	*Cultural*	*LCC*	*Socioec*	*Program*	*Total*
Weights	0.513	0.274	0.074	0.059	0.052	0.028	1

the inconsistency ratio would drop dramatically to 0.0074. This is a very valuable hint pointing to a major source of inconsistency. The matrix shows that the category 'Environment' is 'very strongly' more important than the category 'Cultural' (corresponding entry $a_{12} = 7$) and equally important as 'Life Cycle Cost' ($a_{16} = 1$). Entry $a_{26} = 5$, however, indicates that the category 'Cultural' is 'strongly' more important than 'Life Cycle Cost,' a preference that contradicts the preceding preferences. When the stakeholder was told of this, she attributed it to a trivial error in filling out the questionnaire and asked to change the last entry. As the stakeholder wished to make other changes as well, she was given a new form at the following workshop. Getting more familiar with the questionnaire, the stakeholder improved her inconsistency ratio down to 0.14.

4.2. Extreme inconsistency

Table 8 shows the most inconsistent matrix that was produced (stakeholder no. 7). The stakeholder agreed with the results, but the inconsistency ratio, 0.26, was exceptionally high. This is not surprising, if we note that, while in the second row the categories 'Human Health and Safety' and 'Socioeconomic Impacts' seem to be equally important, each being seven times less important than the category 'Cultural' (i.e., $a_{23} = a_{24} = 7$), their direct comparison records a 9 ($a_{34} = 9$) in favour of 'Human Health and Safety,' whereas a consistent judgement would require $a_{34} = 1$.

A second reason for the high inconsistency ratio of the complete matrix is found in the last two columns: it is evident that the categories 'Environment', 'Cultural,' and 'Human Health and Safety' should be of equal importance. Despite this, we find in the third column that the category 'Cultural' is 'very strongly' more important than HH&S (i.e., $a_{23} = 1/7$), while 'Environment' is 'absolutely' more important than HH&S (i.e., $a_{31} = 1$).

Table 6. Preferences of stakeholder no. 6.

(I)	*HH&S*	*Environ*	*Socioec*	*Cultural*	*Program*	*LCC*	*Total*
Weights	0.333	0.272	0.137	0.114	0.100	0.045	1

(II)	*HH&S*	*Environ*	*Socioec*	*Program*	*Cultural*	*LCC*	*Total*
Weights	0.333	0.272	0.127	0.120	0.104	0.045	1

Table 7. Matrix of pairwise comparisons for SH1.

	Environ	Cultural	HH&S	Socioec	Program	LCC
Environ	1	7	1	5	9	1
Cultural	1/7	1	1/9	1	1	5
HH&S	1	9	1	7	9	9
Socioec	1/5	1	1/7	1	1	1
Program	1/9	1	1/9	1	1	1
LCC	1	1/5	1/9	1	1	1

This is a case of 'scale saturation'. The categories 'Environment', 'Cultural,' and 'Human Health and Safety' are preferred absolutely when compared to categories perceived as much less important ('Programmatic Assumptions' and 'Life Cycle Cost'). This is expressed with the highest possible score ($a_{13} = a_{14} = a_{15} = a_{16} = 9$) resulting in the 'saturation' of the scale, which makes all these categories appear as equally important. However, this does not appear to be true when these 'important categories' are compared among themselves (for instance, $a_{12} = 5$ in lieu of 1), thus leading to inconsistency.

Another surprising result is that this stakeholder ranks the category 'Life Cycle Costs' last, even though he is a representative of the site owner.

A solution could be the use of a wider scale, i.e., not limited to 9. A second solution could be to create two new groups. These would be compared as groups at a new, higher level (see Fig. 4). The two groups of categories would then be compared to each other and the weights composed hierarchically as outlined in Section 2.3.

The discussion of these results among the analysts brought up the general attitude of this stakeholder. There was unanimous agreement that this stakeholder had not taken the whole exercise seriously and that he seemed to be more interested in challenging the analysts than in participating in the process. This was confirmed later, when the stakeholder stopped coming to the meetings, thus depriving us of the opportunity to work with him to revise his assessments along the lines outlined above.

4.3. Whom do the stakeholders represent?

During the elicitation, a question was raised as to whether the stakeholders should represent themselves or their organizations, since their personal views on some matters did not necessarily coincide with those of their organizations. A stakeholder even claimed that she could represent at least three different perspectives: the private citizen's, the employee's, and the employer's. It was decided that a stakeholder should represent his/her institution. Accordingly, it was expected that, if an institution were represented by different representatives at the workshops, the weights obtained from their questionnaires would look very similar. This was confirmed in our case study.

At the second workshop a new person represented the County Environmental Health Department. To ensure consistency, this person was asked to fill out the same AHP questionnaires that his colleague had completed at the first workshop two months

Table 8. Matrix of pairwise comparisons for the most inconsistent questionnaire (SH7).

	Environ	Cultural	HH&S	Socioec	Program	LCC
Environ	1	5	9	9	9	9
Cultural	1/5	1	7	7	9	9
HH&S	1/9	1/7	1	9	9	9
Socioec	1/9	1/7	1/9	1	5	7
Program	1/9	1/9	1/9	1/5	1	1
Cost	1/9	1/9	1/9	1/7	1	1

	Program	LCC	Socioec	Cultural	Environ	HH&S	Total
Weights	0.020	0.018	0.052	0.266	0.507	0.137	1

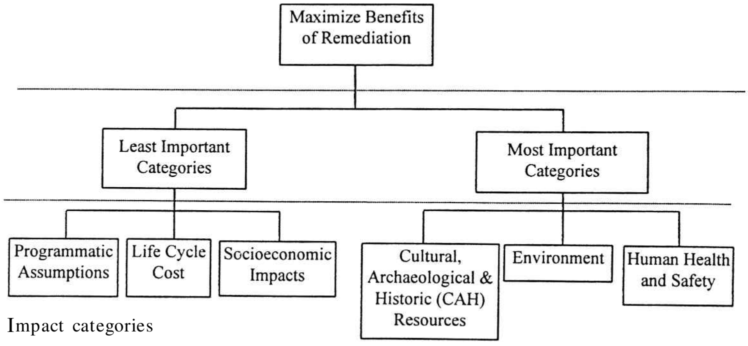

Fig. 4. Additional tier proposed for treatment of extreme inconsistency.

earlier. The first representative was stakeholder no. 11 whose final set of weights is shown in Table 5. Table 9 shows the weights of both stakeholders. Even though the actual weights are different, the overall ranking of the categories by the two stakeholders is the same.

4.4. Results for all the stakeholders

Table 10 shows the weights that our analysis produced for all the stakeholders.

As discussed in Section 2.1, the stakeholders in our case study led us to the development of two value trees (see footnote 1). This is why in Table 10 the weights are reported separately. The top part contains the results for the stakeholders who maintained 'Public Health and Safety' under the category 'Human Health and Safety' (value tree V1), while the bottom part reports the weights of those stakeholders who opted to place 'Public Health and Safety' under the impact category 'Environment' (value tree V2).

It is difficult to identify clear trends. The categories 'Human Health and Safety' and 'Environment' score the highest in both groups, a result that is not unexpected. Stakeholders affiliated with the site owner tended to give higher weight to the category 'Life Cycle Cost' than other stakeholders.

4.5. Sensitivity analysis

The assessed weights were used within the framework of Equation 1 to provide a ranking of the proposed RAAs for each individual stakeholder. The approach used to evaluate the utility preferences of each stakeholder over the impacts of the various performance measures will be presented in a forthcoming paper.

For the validity of the results, an essential phase of the study involves a sensitivity analysis to investigate how the rankings of the alternative remediation actions change when the inputs to the decision analysis differ from the best estimate values. In particular, we have focused our attention on the variability of the subjective importance weight assignment to the various issues of concern in the decision context. The results obtained can be found in Zio and Apostolakis (1998). The analysis has allowed us to test the robustness and stability of the results obtained and provides significant insights which have proven useful for the successive deliberative phase of the process, in which the stakeholders are asked to reach a reasonable degree of consensus on potential recommendations to the regulatory agency on the implementation of a remediation technology.

Table 9. Comparison of rankings for the two representatives.

	HH&S	*Environ*	*Cultural*	*LCC*	*Socioec*	*Program*	*Total*
First repr.							
Weights	0.513	0.274	0.074	0.059	0.052	0.028	1
Second repr.							
Weights	0.398	0.387	0.078	0.070	0.042	0.025	1

5. Conclusions

In this paper we presented, through an application to an environmental management problem, a simple method for eliciting and processing stakeholder priorities. The method provides the weights w_i^k necessary for the calculation of the PI of the different alternatives. In this sense we are not prioritizing the alternatives, yet we are calculating a quantity essential in their prioritization. In particular, AHP was used to perform pairwise comparisons on the impact categories of the top structure of the value tree.

The basic approach to the problem was that no mathematical method could be relied upon to produce accurate weights for the impact categories. Most of the proposed methods in the literature have advantages and disadvantages. Criticisms of the AHP include the possibility of rank reversal and the use of the scale 1–9. These issues have been extensively debated in the literature.

The issue of rank reversal is a major concern in AHP applications. It was first identified by Belton and Gear (1983) and was later discussed by Dyer (1990a, 1990b) and Holder (1990). The question is whether the ranking is preserved when the set of alternatives is changed either by adding or dropping one or more alternatives. A reply was given by Saaty (1990) and Harker and Vargas (1990), with reference to the previous publication by Saaty (1987). Also Forman (1986) provides some insights on the topic. In these papers, it is argued that rank reversal is not always an undesirable effect, and indeed suits several cases. Therefore, it can not be dismissed as a flaw in the algorithm.

In our case the difficulty was overcome by using AHP in its absolute mode. Absolute scales against which the alternatives are compared were built and are the u_{ij}^k of Equation 1. This makes the approach very close to standard Multi-Attribute Utility Theory approaches as both sides of the debate agree, and rules out the possibility of rank reversal.

It should be emphasized that the ranking of the alternatives cannot be achieved by the w_i^k alone. One must use Equation 1, i.e., the utilities must be assessed also (as discussed above). Only then the alternatives for each stakeholder can be ranked. At this point, a deliberation among the stakeholders can begin to reach consensus, as discussed in Apostolakis and Pickett (1998).

We used AHP as a first generator of priorities that had to be reviewed by each stakeholder. We found that the results were generally satisfactory to the stakeholders. We also found that the inconsistency ratio was very valuable to the analysts in discussing the results with the stakeholders and pointing out to them inconsistencies between their pairwise comparisons and their overall rankings of the impact categories.

Table 10. Impact category rankings.

Organization	*Category*	*Programmatic*	*Cost*	*Socioeconomic*	*Cultural*	*Environment*	*Human Health*
V1							
National laboratory employee	SH2	0.022	0.147	0.021	0.059	0.376	0.376
National laboratory contractor	SH5	0.032	0.100	0.042	0.106	0.195	0.524
V2							
Real Estate Agent	SH1	0.078	0.115	0.039	0.042	0.340	0.386
City Environment and Health Department	SH3	0.026	0.070	0.042	0.078	0.387	0.398
Regional Council of Governments	SH4	0.050	0.077	0.249	0.041	0.168	0.415
Community Advisory Board	SH6	0.120	0.045	0.127	0.104	0.272	0.333
Community Involvement and Issues Management	SH7	0.020	0.019	0.052	0.266	0.507	0.137
Native American Nations	SH8	0.035	0.029	0.080	0.155	0.393	0.308
State Environmental Department	SH9	0.033	0.053	0.100	0.026	0.317	0.471
City	SH10	0.196	0.148	0.211	0.06	0.107	0.277
County Environment and Health Department	SH11	0.046	0.045	0.055	0.078	0.276	0.500

In our interactions with the stakeholders, we made sure that computing facilities were available so that the AHP results could be produced in real time as the stakeholders revised their inputs. This was appreciated by the stakeholders and was a significant contributor to building trust in the analysts and the project in general. The more technical stakeholders asked questions about the mathematical foundations of the AHP, while the non-technical stakeholders did not particularly care about the mathematics. They did, however, seem to be impressed by the fact that we were able to point inconsistencies in their inputs, which they also recognized as such and they were willing to modify.

Acknowledgement

This work was sponsored by the United States Department of Energy under cooperative agreement DE-FC01–95EW55088. It is part of a much larger effort that involved risk assessments and stakeholder interactions. We thank our colleagues in this project for making this work possible. They are: Stephie Jennings, Scott Ploger, and Stephanie Weisband of Advanced Sciences, Inc.; Tito Bonano and David Smith of Beta Corporation International; Susan Pickett of MIT; Abbas Ghassemi of New Mexico State University; Pat Salter and Matt Kozak of QuantiSci, Inc.

References

Advanced Sciences, Inc., Beta Corporation International, Massachusetts Institute of Technology, New Mexico State University, and QuantiSci, Inc. (1997), *Risk Communication, Assessment, and Management at Hazardous Waste Sites*, Final Technical Report to the US Department of Energy.

Apostolakis, G.E. and Pickett, S.E. (1998) Deliberation: integrating analytical results into environmental decisions, *Risk Analysis*, accepted for publication.

Belton, V. and Gear, T. (1983) On a short-coming of Saaty's method of analytic hierarchies, *Omega* **11**, 228–30.

Clemen, R.T. (1991) *Making Hard Decisions*, Belmont, California: Duxbury Press.

Dyer, J. S. (1990a) Remarks on the analytic hierarchy process, *Management Science* **36**, 249–58.

Dyer, J. S. (1990b) A clarification of 'Remarks on the Analytic Hierarchy Process', *Management Science* **36**, 274–75.

Edwards, W. and Newman, J.R. (1986) Multiattribute evaluation, in: H.R. Arkes, and K.R. Hammond, (eds), *Judgment and Decision Making: An Interdisciplinary Reader* Cambridge University Press.

Fichtner, J. (1986) On deriving priority vectors from matrices of pairwise comparisons, *Socio-Economic Planning Science*, **20**, 341–45.

Forman, E.H. (1986) Facts and fictions about the analytic hierarchy process, *Mathematical and Computer Modeling* **17**, 4–5.

Gregory, R. and Keeney, R.L. (1994) Creating policy alternatives using stakeholder values, *Management Science* **40**, 1035–48.

Harker, P.T. and Vargas, L.G. (1990) Reply to 'Remarks on the Analytic Hierarchy Process', *Management Science* **36** 269–73.

Holder, R.D. (1990) Some comments on the analytic hierarchy process, *Journal of the Operational Research Society* **41**, 1073–76.

Hong, Y. and Apostolakis, G.E. (1993) Conditional influence diagrams in risk management, *Risk Analysis* **13**, 625–36.

Kadvany, J. (1995) From comparative risk to decision analysis: ranking solutions to multiple-value environmental problems, *Risk* **6**, 333–58.

Keeney, R. (1981) Analysis of preference dependencies among objectives, *Operations Research* **29**, 1105–20.

Keeney, R.L. and von Winterfeldt, D. (1994) Managing nuclear waste from power plants, *Risk Analysis* **14**, 107–30.

Merkhofer, M.W. and Keeney, R.L. (1987) A multiattribute utility analysis of alternative sites for the disposal of nuclear waste, *Risk Analysis* **7**, 173–94.

National Research Council (1996) *Understanding Risk: Informing Decisions in a Democratic Society,* Washington, DC: National Academy Press.

Saaty, T.L. (1987) Rank generation, preservation, and reversal in the analytic hierarchy process, *Decision Science* **18**, 157–77.

Saaty, T.L. (1990) An exposition of AHP in reply to the paper 'Remarks on the Analytic Hierarchy Process', *Management Science* **36**, 259–68.

Saaty, T. L. (1996) *The Analytic Hierarchy Process*, Pittsburgh: RWS Publications.

Zio, E. and Apostolakis, G.E. (1998) Application of sensitivity analysis to environmental management, in *Proceedings of the Second International Symposium on Sensitivity Analysis of Model Output*, SAMO98, Venice, April, pp. 333–36.